Near Heisenberg limited atomic clocks in the presence of decoherence

J. Borregaard[1] and A. S. Sørensen[1]

[1]*QUANTOP, Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen , Denmark*

(Dated: April 9, 2013)

The stability of current atomic clocks is limited by the quantum noise of the atoms. To reduce this noise it has been suggested to use entangled atomic ensembles with reduced atomic noise. Potentially this can push the stability all the way to the limit allowed by the Heisenberg uncertainty relation, which is denoted as the Heisenberg limit. In practice, however, entangled states are often more prone to decoherence, which may prevent reaching this performance. Here we present an adaptive measurement protocol that in the presence of a realistic source of decoherence enables us to get near Heisenberg limited stability of atomic clocks using entangled atoms. The protocol may realize the full potential of entanglement for quantum metrology despite the detrimental influence of decoherence.

PACS numbers: 06.30.Ft, 03.65.Yz, 03.65.Ud, 06.20.Dk

Some of the most precise measurements in physics are made with atomic clocks. Very accurate time measurements are used in anything from GPS-systems to gravitational wave detectors. Currently one of the main limitations to the stability of atomic clocks comes from the quantum noise of the atoms, which leads to the standard quantum limit (SQL) where the resolution scales as $1/\sqrt{N}$ with N being the number of atoms [1, 2]. To overcome this noise it has been suggested to use entangled states with reduced atomic noise [3–7]. Ultimately this may lead to a stability at the Heisenberg limit where the resolution scales as $1/N$, and recently the first proof of principle experiments have demonstrated these concepts experimentally [8–13]. In practice, however, entangled states are often more prone to decoherence, and to fully asses the advantage it is essential to study the performance in the presence of decoherence [14]. In Ref. [5] it was proven that entanglement can be used to improve the long-term stability of atomic clocks in the presence of the dominant practical source of decoherence, but the improvement identified was rather limited. Here we show that it is possible to obtain a large improvement in the resolution of the atomic clock by combining entanglement with an adaptive measurement protocol (inspired by Ref. [15]). With our adaptive measurement protocol the entangled states are not more sensitive to the decoherence than disentangled states. As a consequence the long term stability of the atomic clock can be improved almost to the Heisenberg limit even in the presence of decoherence.

An atomic clock operates by locking a local oscillator (LO) to an atomic transition via a feedback loop. The feedback is based on a measurement of the LO frequency offset $\delta\omega_0$ compared to the atomic transition. The frequency offset is typically obtained through Ramsey spectroscopy [16]. Ramsey spectroscopy consists of three parts; first the atoms are illuminated by a near-resonant $\pi/2$-pulse from the LO followed by the Ramsey time T of free evolution, and finally another near-resonant $\pi/2$-pulse is applied to the atoms. During the free evolution the LO acquires a phase $\delta\phi_0 = \delta\omega_0 T$ relative to the atoms. Due to the $\pi/2$-pulses this phase can be measured as a population difference between the two clock levels. $\delta\omega_0$ can thus be estimated from the measurement and used for a feedback that steers the frequency of the LO to the clock-frequency. The resolution of the clock will improve with T since a longer T improves the relative sensitivity of the phase measurement. For current atomic fountain clocks, T is limited by gravity and can hardly be varied. Here on the other hand we consider trapped particles, where T can be increased until it is limited by the decoherence in the system. The long term stability thus depends on the nature of the decoherence. To take this into account Ref. [14] considered dephasing of single atoms to be the dominant noise in the system. For this model Ref. [14] showed that entanglement can not improve the stability of atomic clocks considerable (although an improvement is possible for non-Markovian noise [17]). A more realistic model of the decoherence was described in Ref. [5] where the primary noise source is the frequency fluctuations of the LO [18]. In this work a small improvement in the long term stability, scaling as $\sim N^{1/6}$, was identified for entangled atoms. Here we use the same realistic model of the LO noise and show that entanglement and adaptive measurements may improve the performance and give near Heisenberg limited atomic clocks.

We consider an ensemble of N two-level atoms, which we model as a collection of spin-1/2 particles. The total angular momentum is $\vec{J} = \sum_{i=1}^{N} \vec{S}_i$ where $\vec{S}_i$ is the spin-vector of the i'th atom. The angular momentum operators $\hat{J}_{x,y,z}$ gives the projections of $\vec{J}$ on the x, y and z-axis. The atoms are initially pumped to have a mean spin along the z-axis, $\langle \hat{J}_x \rangle = \langle \hat{J}_y \rangle = 0$. After the Ramsey sequence the Heisenberg evolution of $\hat{J}_x, \hat{J}_y$ and $\hat{J}_z$ is $\hat{J}_1 = \hat{J}_x, \hat{J}_2 = \sin(\delta\phi_0)\hat{J}_y - \cos(\delta\phi_0)\hat{J}_z$ and

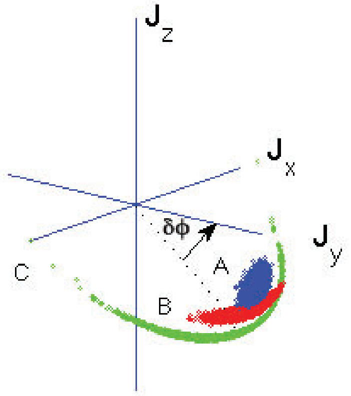

FIG. 1. (Colour online) The atomic state just before the measurement of J_z for (A) uncorrelated atoms, (B) moderate squeezed atoms and (C) highly squeezed atoms. Highly squeezed atoms introduce significant extra noise in the measurement through a noise term $\propto \delta\phi\Delta\hat{J}_z/|\langle J_z\rangle|$ in the phase estimate. The adaptive protocol seeks to rotate the atomic state to have mean spin almost along the J_y-axis thus reducing this noise.

$\hat{J}_3 = \cos(\delta\phi_0)\hat{J}_y + \sin(\delta\phi_0)\hat{J}_z$. At the end of the Ramsey sequence $\hat{J}_3$ is measured and used to estimate $\delta\phi_0$. The $\hat{J}_y$ term in $\hat{J}_3$ results in the so called projection noise in the phase estimate $\sim \Delta J_y/|\langle J_z\rangle|$. For uncorrelated atoms $\Delta\hat{J}_y\Delta\hat{J}_x = \langle\hat{J}_z\rangle/2 \approx N/4$ and the projection noise causes the stability of the clock to scale as $\sim 1/\sqrt{N}$. For a spin squeezed state [19] the variance of $\hat{J}_y$ is reduced to obtain a better phase estimate. Such a spin squeezed state is depicted in Fig. 1, which shows how the noise of the spin squeezed state looks like a "flat banana" lying on the Bloch sphere. The more we squeeze, the longer and more narrow the banana is and significant extra noise is added to the mean spin direction. For a phase measurement based on a direct measurement of $\hat{J}_3$ this gives an additional noise term $\sim \delta\phi_0\Delta\hat{J}_z/|\langle J_z\rangle|$ in the phase estimate. This extra noise limited the performance in Ref. [5] if strongly squeezed states were used such that only a limited improvement could be identified. We avoid this problem by using an adaptive scheme with weak measurements to make a rough estimate of $\delta\phi_0$ and then rotate the spins of the atoms such that the mean spin direction is almost along the y-axis. The flat banana depicted in Fig. 1 will afterwards lie in the xy-plane and this will eliminate any noise from $\Delta\hat{J}_z$ in subsequent measurements. Having eliminated the noise from $\Delta\hat{J}_z$ we can allow strong squeezing in $\Delta\hat{J}_y$ and obtain near Heisenberg limited stability.

To be concrete our weak measurements is based on the strategy developed and demonstrated in Ref [10, 11, 20] where a light field dispersively interact with the spin and is subsequently measured. This is described by a Hamiltonian $H_{int} = -\chi_1\hat{J}_3\hat{X}_1$ where χ_1 is the interactions strength and $\hat{X}_1$ is the canonical position operator of the light [21–23]. The measurement results in a rotation around $\hat{J}_3$ described by the rotation matrix $\mathbf{R}_3(\hat{\Pi}_1)$, where $\hat{\Pi} = \Omega_1\hat{X}_1$. $\Omega_1 = \chi_1\mu_1$ is the measurement strength and μ_1 is the measurement time. The canonical momentum operators of the light before $\hat{P}_1$ and after $\hat{P}_1'$ the interaction are then related by $\hat{P}_1' = \hat{P}_1 - \Omega_1\hat{J}_3$. $\hat{P}_1'$ is measured using homodyne detection [24] and the phase is estimated as $\delta\phi_1^e = \frac{-\beta_1\hat{P}_1'}{\Omega_1\langle\hat{J}_z\rangle}$ where the factor β_1 is found from minimizing $\langle(\delta\phi_0 - \delta\phi_1^e)^2\rangle$. Based on the phase estimate we rotate the spin of the atoms around $\hat{J}_1$ in order to compensate for the extra noise added by the spin squeezing. This is described by a rotation matrix $\mathbf{R}_1(\delta\phi_1^e)$. The process can be iterated such that after $n-1$ weak measurements the Heisenberg evolution of the original operators $(\hat{J}_1, \hat{J}_2, \hat{J}_3)$ is:

$$\begin{pmatrix}\hat{J}_1\\ \hat{J}_2\\ \hat{J}_3\end{pmatrix}_n = \mathbf{R}_1(\delta\phi_{n-1}^e)\mathbf{R}_3(\hat{\Pi}_{n-1})...\mathbf{R}_1(\delta\phi_1^e)\mathbf{R}_3(\hat{\Pi}_1)\begin{pmatrix}\hat{J}_1\\ \hat{J}_2\\ \hat{J}_3\end{pmatrix} \quad (1)$$

The final measurement is assumed to be a projective measurement and the final phase estimate $\delta\phi_n^e$ is thus $\delta\phi_n^e = \frac{\beta_n\hat{J}_{3,n}}{\langle\hat{J}_z\rangle}$. The factors of β_i in the phase estimates are found be minimizing $\langle(\delta\phi_0 - \sum_{i=1}^{i}\delta\phi_i^e)^2\rangle$ with respect to β_i after each measurement.

We will now show semi-analytically that the measurement strategy in equation (1) allows for near Heisenberg limited stability. For simplicity we set all $\beta_i = 1$. After a series of n adaptive measurements the estimate $\delta\phi^e$ of the initial phase of the atoms $(\delta\phi_0)$ is $\delta\phi^e = \sum_{i=1}^{n}\delta\phi_i^e$. $\delta\phi^e$ is then used for the feedback on the LO. We assume that negligible time is spent on everything but the free evolution period T. After a time $t_k = kT$ the frequency correction $\Delta\omega(t_k) = -\alpha\delta\phi_e(t_k)/T$ is applied to the LO where α sets the strength of the feedback loop. The frequency offset of the LO at time t_k is then $\delta\omega(t_k) = \delta\omega_0(t_k) + \sum_{i=1}^{k}\Delta\omega(t_i)$, where $\delta\omega_0(t_k)$ is the frequency fluctuation of the unlocked LO and $\Delta\omega(t_i)$ is the frequency correction applied at time t_i. The mean frequency offset after running for a period $\tau = lT$ $(l \gg 1)$ is

$$\delta\bar{\omega}(\tau) = \frac{1}{l}\sum_{k=1}^{l}\frac{\delta\phi(t_k) - \delta\phi_e(t_k)}{T}, \quad (2)$$

where $\delta\phi(t_k)$ is the accumulated phase of the LO during the free evolution between t_{k-1} and t_k and $\delta\phi_e(t_k)$ is the estimate of this phase (see supplementary information). The measure of the long term stability of the atomic clock is:

$$\sigma_\gamma(\tau) = \langle(\delta\bar{\omega}(\tau)/\omega)^2\rangle^{1/2} \quad (3)$$

$$= \sqrt{\frac{1}{\tau\omega^2}}\left(\frac{1}{l}\frac{\langle(\sum_{k=1}^{l}\delta\phi(t_k) - \delta\phi_e(t_k))^2\rangle}{T}\right)^{1/2} \quad (4)$$

We initially assume that the phase offset of the unlocked LO $\delta\phi_0$ is due to phase fluctuations in the LO with a white noise spectrum. Later we will also consider the case where the fluctuations have a $1/f$ spectrum. For white noise we have $\langle\delta\phi_0^2\rangle = \gamma T$ ($\langle\delta\phi_0\rangle = 0$) where γ is a parameter characterizing the fluctuations of the LO. In the limit of $\alpha \ll 1$, the phases are uncorrelated such that $\sigma_\gamma(\tau) = \sqrt{\gamma/\tau\omega^2}(\langle(\delta\phi_0 - \delta\phi^e)^2\rangle/\gamma T)^{1/2}$. This expression shows that for fixed γ, τ the stability of the clock only depends on how precisely we can estimate $\delta\phi_0$.

We now turn to the adaptive measurement strategy to estimate $\delta\phi_0$. After j weak measurements the difference between the true phase and the estimated phase $\delta\Phi_j$ is:

$$\delta\Phi_j = \delta\phi_0 - \sum_{i=1}^{j}\delta\phi_i^e = \delta\phi_0 - \sum_{i=1}^{j-1}\delta\phi_i^e - \delta\phi_j^e. \quad (5)$$

Using equation (1) to get an expression for $\delta\phi_j^e$ and the fact that $\delta\Phi_{j-1} = \delta\phi_0 - \sum_{i=1}^{j-1}\delta\phi_i^e$, we can express the phase error as

$$\delta\Phi_j \approx \delta\Phi_{j-1}(1 - \hat{J}_z/\langle\hat{J}_z\rangle) - (\hat{J}_y + \delta\hat{J}_{3,j} - \hat{P}_j/\Omega_j)/\langle\hat{J}_z\rangle, (6)$$

where we have assumed $\delta\Phi_{j-1} \ll 1$. The first term in (6) gives a contribution $\sim \delta\Phi_{j-1}\Delta J_z/\langle\hat{J}_z\rangle$ to $\sigma_\gamma(\tau)$ from the noise in the mean spin direction as previously mentioned. Note that this term is proportional to the phase estimation error at the previous measurement stage, since it depends on how well the banana in Fig. 1 is rotated into the xy-plane. For a useful adaptive protocol $\delta\Phi_{j-1}$ gets smaller for growing j and the noise that enters through $\Delta\hat{J}_z$ is reduced. The last terms in (6) gives the noise from $\Delta\hat{J}_y$, the accumulated back action of the previous measurements ($\langle\delta J_{3,j}^2\rangle$) and the noise from the incoming light in the measurement ($\Delta\hat{P}_j^2 = \langle\hat{P}_j^2\rangle$).

The stronger a measurement is, the less noise is added through $\Delta\hat{P}_j^2/\Omega_j^2$ since the measurement is more precise. Any imprecision $\Delta\hat{P}_{i<j}^2/\Omega_{i<j}^2$ from previous measurements is contained in $\delta\Phi_{j-1}$ and is corrected for in the subsequent stages of the protocol, which essentially estimate how well we corrected the phase in previous measurements. This means that we can initially work with weak measurements, which only give a rough estimate since later stronger measurements correct for the imprecision in the initial measurements.

The accumulated back action noise $\delta\hat{J}_{3,j}$ originates from the disturbance caused be the measurements. The measurements adds noise in $\hat{J}_1, \hat{J}_2$, which is mixed into $\hat{J}_3$ when the atomic state is rotated to have mean spin almost along the y-axis. From equation (1) the dominant terms in $\delta\hat{J}_{3,j}$ is found to be $\delta\hat{J}_{3,j} \sim \sum_{i=1}^{j-1}\delta\phi_i^e\Omega_i\hat{X}_i\hat{J}_x$ (see supplementary information). It is seen that the stronger a measurement is, the more noise is added to the stability. However for a useful adaptive protocol $\delta\phi_i^e$ gets smaller for growing i, which again means that the i'th measurement can be stronger than the previous $(i-1)$'th measurement without adding more noise to the stability.

Above we have argued that we can suppress the noise terms originating from ΔJ_z, $\Delta\hat{P}_j$, and $\delta\hat{J}_{3,j}$ using an adaptive protocol with weak initial measurements. A remaining question is how well this suppression work. This is considered in detail in the supplementary information. To be specific, we consider spin squeezed states of the form $|\psi(\kappa)\rangle = \mathcal{N}(\kappa)\sum_m(-1)^m e^{-(m/\kappa)^2}|m\rangle$, where $|m\rangle$ are eigenstates of $\hat{J}_y$ with eigenvalue m, $\mathcal{N}(\kappa)$ is a normalization constant and the sum is from $-J$ to J where $J = N/2$ is the total angular momentum quantum number. κ gives the width of the state and thus determines the degree of squeezing. Uncorrelated atoms are well approximated by $\kappa = N^{1/2}$ while highly squeezed states are obtained for $\kappa \to 1$. We furthermore assume that the probe light has vacuum statistics such that $\langle\hat{P}\rangle = \langle\hat{X}\rangle = \langle\hat{X}\hat{P}\rangle = 0$ and $\langle\hat{X}^2\rangle = \langle\hat{P}^2\rangle = 1/2$. As an upper limit of the stability we find that for $n \sim 3\log(N)$ weak measurements, using a spin squeezed state with $\kappa \sim \log\sqrt{N} + 2$ and choosing a measurement strategy with $\Omega_i \sim N^{-1+i/(n+1)}$ we can suppress other noise terms so that the measurement is eventually limited by $\Delta\hat{J}_y$. $\sigma_\gamma(\tau)$ will then be $\sim (2/N + \log\sqrt{N}/N)/\sqrt{\gamma T}$ (in units of $\sqrt{\gamma/\tau\omega^2}$) for $N \gg 1$. It is seen that for $N = 10^6$ the upper limit of the stability will differ from the Heisenberg limit, which is $\sigma_\gamma(\tau) = (1/N)/\sqrt{\gamma T}$ by a factor of ~ 5.

The above upper limit shows that we can be near the Heisenberg limit, but to get a better estimate we have made a numerical minimization of σ_γ in order to find the optimal resolution. We have simulated an atomic clock with a LO subject to white Gaussian noise for atom numbers in the range $N = 100$ to $N = 10^6$ (for details see supplementary information). For $N \leq 1000$ we simulate the full quantum evolution (we denote this as 'full quantum simulation') and for $N > 1000$ we approximate the probability distributions of $\hat{J}_{x,y,z}$ with Gaussian distributions with moments calculated from $|\psi(\kappa)\rangle$ in the approximation $N \gg 1$ (we denote this as 'Gaussian simulation').

An example of the results are shown in Fig. 2, which demonstrates that there is a significant improvement by using a spin squeezed state compared to uncorrelated atoms. Also with the adaptive protocol, the Ramsay time can be as large for highly entangled states as for disentangled state and there is thus no difference in the relevant coherence time. Furthermore the adaptive protocol allows longer interrogation times $\gamma T \lesssim 0.3$ than the conventional protocol $\gamma T \lesssim 0.1$. This is because the adaptive protocol can determine phases $\lesssim \pi$ while the conventional protocol begin to give ambiguous results for phases $\sim \pi/2$.

We have numerically minimized $\sigma_\gamma(\tau)$ in both the degree of squeezing, the number of weak measurements, the Ramsey time, and the strengths of the measure-

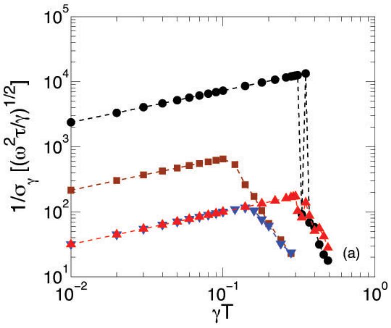

FIG. 2. (Colour online) Stability as a function of the Ramsey time (γT) for $N = 10^5$. $\square, \triangledown$ are the best non-linear protocol of Ref. [5] while $\circ, \triangle$ are the adaptive protocol discussed in this article. The adaptive protocol allows for $\gamma T \sim 0.3$ while the conventional protocol of Ref. [5] only allows for $\gamma T \sim 0.1$. $\triangledown, \triangle$ correspond to uncorrelated atoms while $\circ, \square$ are the ideal choices of squeezing in the adaptive and conventional protocols respectively. For the adaptive protocol the figure shows how the probabilistic nature of the errors that limits γT makes the stability jump back and forth for $\gamma T > 0.3$ before it is definitely diminished (see supplementary information).

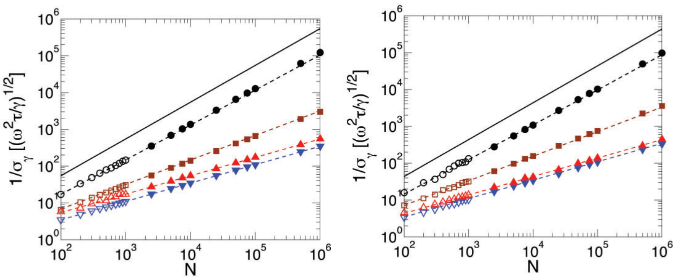

FIG. 3. (Colour online) Optimized stability of an atomic clock for a LO subject to (a) white noise and (b) $1/f$-noise. $\circ, \square, \triangle, \triangledown$ are the results for the full quantum simulation while $\bullet, \blacksquare, \blacktriangle, \blacktriangledown$ are the results for the Gaussian simulation. The Gaussian simulation has been extended down to $N = 100$ and was found to give more or less identical results to the full quantum simulation. $\circ,\bullet$ ($\blacktriangle,\triangle$) are the adaptive scheme and $\square$, $\blacksquare$ ($\triangledown,\blacktriangledown$) are for the conventional protocol with (without) entanglement. The dotted lines are the analytical results and the solid line is the Heisenberg limit for the maximal Ramsay time $\gamma T = 0.3$ (a) and $\gamma T = 0.2$ (b).

ments. Fig. 3(a) shows the result of the optimization for both the adaptive protocol and the conventional protocol with/without squeezing. The adaptive protocol gives a significant improvement compared to using uncorrelated atoms resulting in near Heisenberg limited stability. The numerical calculations also agree nicely with the analytical calculations (see supplementary information for details). As noted above the adaptive protocol allow for a longer Ramsey time than the conventional protocol, which gives an improvement of roughly a factor 1.6 for uncorrelated atoms (see Fig. 2). Using the adaptive measurement protocol thus leads to an improvement even without entanglement.

So far we have assumed white noise in the LO. In practice the noise of the LO is however more likely to have a nontrivial spectrum like $1/f$-noise. We have therefore repeated the numerical simulations with the LO being subject to $1/f$-noise for which $\langle \omega(f)\omega(f') \rangle = \delta(f + f')\gamma^2/f$. The details of the locking of the LO to the atoms through a feedback is described in the supplementary information. We have performed the numerical optimization in the same parameters as for white noise. The results are shown in Fig. 3(b) and shows that the improvement obtained using the adaptive scheme with correlated atoms persists also for $1/f$ noise. Again near Heisenberg limited stability is obtained using the adaptive protocol. The longer Ramsey time of the adaptive scheme compared to projective measurements gives an improvement of roughly a factor 1.3 for uncorrelated atoms.

In conclusion we have developed an adaptive measurement protocol which allows operating atomic clocks near the Heisenberg limit using entangled spin squeezed ensembles of atoms. These results clearly demonstrates that entanglement can be an important resource for quantum metrology. Importantly our results are obtained under realistic assumptions where we account for the dominant source of noise in practice. We find that in this situation we can gain nearly the full potential of entanglement estimated without accounting for decoherence. Furthermore the adaptive protocol allows for a higher Ramsey time, which gives an improvement even for uncorrelated atoms.

We gratefully acknowledge the support of the Lundbeck Foundation, QUANTOP and the Danish National Research Foundation. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 306576.

[1] G. Santarelli, P. Laurent, P. Lemonde, A. Clairon, A. G. Mann, S. Chang, A. N. Luiten, and C. Salomon, Phys. Rev. Lett. **82**, 4619 (1999).

[2] W. M. Itano, J. C. Bergquist, J. J. Bollinger, J. M. Gilligan, D. J. Heinzen, F. L. Moore, M. G. Raizen, and D. J. Wineland, Phys Rev. A **47**, 3554 (1993).

[3] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen, Phys Rev. A **50**, 67 (1994).

[4] J. J. Bollinger, W. M. Itano, D. J. Wineland, and D. J. Heinzen, Phys Rev. A **54**, R4649 (1996).

[5] A. André, A. S. Sørensen, and M. D. Lukin, Phys. Rev. Lett. **92**, 230801 (2004).

[6] T. Rosenband, arXiv:1203.0288v2 [quant-ph] (2012).

[7] V. Giovanetti, S. Lloyd, and L. Maccone, nature photonics **5** (2011), 10.1038/nphoton.2011.35.

[8] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton,

J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland, Science **304**, 1476 (2004).

[9] I. D. Leroux, M. H. Schleier-Smith, and V. Vuletić, Phys. Rev. Lett. **104**, 073602 (2010).

[10] J. Appel, P. J. Windpassinger, D. Oblak, U. B. Hoff, N. Kjærgaard, and E. S. Polzik, Proceedings of the National Academy of Sciences **106**, 10960 (2009).

[11] A. Louchet-Chauvet, J. Appel, J. J. Renema, D. Oblak, N. Kjaergaard, and E. S. Polzik, New Journal of Physics **12**, 065032 (2010).

[12] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, and P. Treutlein, Nature **464**, 1170 (2010).

[13] C. Gross, T. Zibold, E. Nicklas, J. Esteve, and M. K. Oberthaler, Nature **464**, 1165 (2010).

[14] S. F. Huelga, C. Macchiavello, T. Pellizzari, A. K. Ekert, M. B. Plenio, and J. I. Cirac, Phys. Rev. Lett. **79**, 3865 (1997).

[15] D. W. Berry and H. M. Wiseman, Phys. Rev. Lett. **85**, 5098 (2000).

[16] N. F. Ramsey, *Molecular Beams* (Oxford University Press, London, 1956).

[17] A. W. Chin, S. F. Huelga, and M. B. Plenio, Phys. Rev. Lett. **109**, 233601 (2012).

[18] D. J. Wineland, C. Monroe, W. M. Itano, D. Leibfried, B. E. King, and D. M. Meekhof, J. Res. Natl. Inst. Stand. Technol. **103**, 259 (1998).

[19] M. Kitagawa and M. Ueda, Phys. Rev. A **47**, 5138 (1993).

[20] J. L. Sørensen, J. Hald, and E. S. Polzik, Phys. Rev. Lett. **80**, 13487 (1998).

[21] K. Hammerer, K. Mølmer, E. S. Polzik, and J. I. Cirac, Phys. Rev. A **70**, 044304 (2004).

[22] L.-M. Duan, J. I. Cirac, P. Zoller, and E. S. Polzik, Phys. Rev. Lett. **85**, 5643 (2000).

[23] A. Kuzmich, L. Mandel, and N. P. Bigelow, Phys. Rev. Lett. **85**, 1594 (2000).

[24] K. Hammerer, A. S. Sørensen, and E. S. Polzik, Rev. Mod. Phys. **82**, 1041 (2010).

[25] A. André, *Nonclassical states of light and atomic ensembles: Generation and New Applications*, Ph.D. thesis, Harvard Univeristy, Cambridge Massachusetts (2005).

SUPPLEMENTARY INFORMATION

This supplementary material to our article "Near Heisenberg limited atomic clocks in the presence of decoherence" consists of two parts. In the first part we will go through the adaptive protocol in detail and describe the various noise terms that effect the stability of the clock. We show semi-analytically that the adaptive protocol gives near Heisenberg limited clocks with a stability scaling at most like $\log N/N$. The second part describes the details of our numerical simulations of an atomic clock. We describe the subtleties of simulating an atomic clock running for a long but finite time and show how the feedback on the LO effectively locks it to the atoms.

SCALING OF THE STABILITY

In the article we describe how the adaptive protocol consists of a series of weak measurements of the atomic spin with intermediate feedback on the atoms. After a weak measurement we estimate the phase of the LO relative to the atoms as $\delta\phi_i^e = \frac{-\beta_i \hat{P}_i'}{\Omega_i \langle \hat{J}_z \rangle}$ where $\hat{P}_i'$ is the canonical momentum operator of the light after the interaction with the atoms and Ω_i is the measurement strength. To simplify the equations we set $\beta_i = 1$ throughout this section. Note that this choice is not ideal and the true performance of the clock will thus be better than what we estimate here. In order to go through the details of the adaptive protocol we consider a Ramsey sequence where the LO acquires a phase $\delta\phi_0$ relative to the atoms during the free evolution. $\delta\phi_0$ is subsequently estimated through $n-1$ weak measurements with intermediate feedback on the atoms and a final projective measurement. We define the operators

$$\hat{J}_1 = \hat{J}_x, \qquad \hat{J}_2(\theta) = \sin(\theta)\hat{J}_y - \cos(\theta)\hat{J}_z, \qquad \hat{J}_3(\theta) = \sin(\theta)\hat{J}_z + \cos(\theta)\hat{J}_y \tag{7}$$

such that the outcome of the Heisenberg evolution of $\hat{J}_x, \hat{J}_y$ and $\hat{J}_z$ during the Ramsey sequence is $\hat{J}_1, \hat{J}_2(\delta\phi_0)$ and $\hat{J}_3(\delta\phi_0)$. In the Heisenberg picture the first weak measurement results in a rotation of the atomic spin described by a rotation matrix $\mathbf{R}_3(\hat{\Pi}_1)$, where $\hat{\Pi} = \Omega_1 \hat{X}_1$ as described in the article ($\hat{X}_1$ is the canonical position operator of the probe light). The subsequent feedback on the atoms is a rotation described by the rotation matrix $\mathbf{R}_1(\delta\phi_1^e)$ where $\delta\phi_1^e$ is the phase estimate based on the measurement result. Thus the Heisenberg evolution of the operators after the first weak measurement and subsequent feedback is:

$$\begin{pmatrix} \hat{J}_1 \\ \hat{J}_2 \\ \hat{J}_3 \end{pmatrix}_2 = \mathbf{R}_1(\delta\phi_1^e)\mathbf{R}_3(\hat{\Pi}_1) \begin{pmatrix} \hat{J}_1 \\ \hat{J}_2(\delta\phi_0) \\ \hat{J}_3(\delta\phi_0) \end{pmatrix} \tag{8}$$

$$= \begin{pmatrix} \cos\hat{\Pi}_1 & -\sin\hat{\Pi}_1 & 0 \\ \cos\delta\phi_1^e \sin(\hat{\Pi}_1) & \cos(\delta\phi_1^e)\cos(\hat{\Pi}_1) & -\sin(\delta\phi_1^e) \\ \sin(\delta\phi_1^e)\sin(\hat{\Pi}_1) & \sin(\delta\phi_1^e)\cos(\hat{\Pi}_1) & \cos(\delta\phi_1^e) \end{pmatrix} \begin{pmatrix} \hat{J}_1 \\ \hat{J}_2(\delta\phi_0) \\ \hat{J}_3(\delta\phi_0) \end{pmatrix} \tag{9}$$

$$= \begin{pmatrix} \hat{J}_1 \\ \hat{J}_2(\delta\phi_0 - \delta\phi_1^e) \\ \hat{J}_3(\delta\phi_0 - \delta\phi_1^e) \end{pmatrix} + \begin{pmatrix} \delta\hat{J}_{1,2} \\ \delta\hat{J}_{2,2} \\ \delta\hat{J}_{3,2} \end{pmatrix} \tag{10}$$

where the operators

$$\delta\hat{J}_{1,2} = (\cos(\hat{\Pi}_1) - 1)\hat{J}_1 - \sin(\hat{\Pi}_1)\hat{J}_2(\delta\phi_0) \tag{11}$$

$$\delta\hat{J}_{2,2} = \cos(\delta\phi_1^e)\sin(\hat{\Pi}_1)\hat{J}_1 + \cos(\delta\phi_1^e)(\cos(\hat{\Pi}_1) - 1)\hat{J}_2(\delta\phi_0) \tag{12}$$

$$\delta\hat{J}_{3,2} = \sin(\delta\phi_1^e)\sin(\hat{\Pi}_1)\hat{J}_1 + \sin(\delta\phi_1^e)(\cos(\hat{\Pi}_1) - 1)\hat{J}_2(\delta\phi_0) \tag{13}$$

describe the noise due to the back action of the measurement. Note that the back action also effects $\hat{J}_3$ even though $\hat{J}_3$ is conserved and hence unaffected during the measurement. This is because the rotation during the feedback mixes back action noise into $\hat{J}_3$.

The process is now iterated such that the Heisenberg evolution after $n-1$ weak measurements and subsequent rotations is

$$\begin{pmatrix} \hat{J}_1 \\ \hat{J}_2 \\ \hat{J}_3 \end{pmatrix}_n = \begin{pmatrix} \hat{J}_1 \\ \hat{J}_2(\delta\phi_0 - \sum_{i=1}^{n-1} \delta\phi_i^e) \\ \hat{J}_3(\delta\phi_0 - \sum_{i=1}^{n-1} \delta\phi_i^e) \end{pmatrix} + \begin{pmatrix} \delta\hat{J}_{1,n} \\ \delta\hat{J}_{2,n} \\ \delta\hat{J}_{3,n} \end{pmatrix} \tag{14}$$

where the iterative expressions of $\delta\hat{J}_{1,j-1}, \delta\hat{J}_{2,j-1}, \delta\hat{J}_{3,j-1}$ are

$$\delta\hat{J}_{1,j} = (\cos(\hat{\Pi}_{j-1})-1)\hat{J}_1 - \sin(\hat{\Pi}_{j-1})\hat{J}_2\left(\delta\phi_0 - \sum_{i=1}^{j-1}\delta\phi_i^e\right) + \cos(\hat{\Pi}_{j-1})\delta\hat{J}_{1,j-1} - \sin(\hat{\Pi}_{j-1})\delta\hat{J}_{2,j-1} \tag{15}$$

$$\delta\hat{J}_{2,j} = \cos(\delta\phi_{j-1}^e)\sin(\hat{\Pi}_{j-1})\hat{J}_1 + \cos(\delta\phi_{j-1}^e)(\cos(\hat{\Pi}_{j-1})-1)\hat{J}_2\left(\delta\phi_0 - \sum_{i=1}^{j-1}\delta\phi_i^e\right) + \cos(\delta\phi_{j-1}^e)\sin(\hat{\Pi}_{j-1})\delta\hat{J}_{1,j-1} \tag{16}$$

$$+ \cos(\delta\phi_{j-1}^e)\cos(\hat{\Pi}_{j-1})\delta\hat{J}_{2,j-1} - \sin(\delta\phi_{j-1}^e)\delta\hat{J}_{3,j-1}$$

$$\delta\hat{J}_{3,j} = \sin(\delta\phi_{j-1}^e)\sin(\hat{\Pi}_{j-1})\hat{J}_1 + \sin(\delta\phi_{j-1}^e)(\cos(\hat{\Pi}_{j-1})-1)\hat{J}_2\left(\delta\phi_0 - \sum_{i=1}^{j-1}\delta\phi_i^e\right) + \sin(\delta\phi_{j-1}^e)\sin(\hat{\Pi}_{j-1})\delta\hat{J}_{1,j-1}$$

$$+ \sin(\delta\phi_{j-1}^e)\cos(\hat{\Pi}_{j-1})\delta\hat{J}_{2,j-1} + \cos(\delta\phi_{j-1}^e)\delta\hat{J}_{3,j-1} \tag{17}$$

with $\delta\hat{J}_{1,1} = \delta\hat{J}_{2,1} = \delta\hat{J}_{3,1} = 0$. In the final *projective* measurement we measure $\hat{J}_{3,n}$ and obtain a phase estimate $\delta\phi_n^e = \hat{J}_{3,n}/\langle\hat{J}_z\rangle$. Our final estimate of $\delta\phi_0$ is then $\sum_{i=1}^{n}\delta\phi_i^e$ and the difference between this and the true phase is

$$\delta\Phi_n = \delta\phi_0 - \sum_{i=1}^{n}\delta\phi_i^e = \delta\phi_0 - \sum_{i=1}^{n-1}\delta\phi_i^e - \delta\phi_n^e. \tag{18}$$

Using equation (14) we can express this using the previous phase estimation error ($\delta\Phi_{n-1}$) and the last measurement

$$\delta\Phi_n = \delta\Phi_{n-1} - \left(\sin(\delta\Phi_{n-1})\hat{J}_z + \cos(\delta\Phi_{n-1})\hat{J}_y + \delta\hat{J}_{3,n}\right)/\langle\hat{J}_z\rangle \tag{19}$$

$$\approx \delta\Phi_{n-1}(1 - \hat{J}_z/\langle\hat{J}_z\rangle) - \hat{J}_y/\langle\hat{J}_z\rangle - \delta\hat{J}_{3,n}/\langle\hat{J}_z\rangle \tag{20}$$

where $\delta\Phi_{n-1} = \delta\phi_0 - \sum_{i=1}^{n-1}\delta\phi_i^e$ and we have assumed $\delta\Phi_{n-1} \ll 1$ to expand the sine and cosine. Below we will argue that for white noise and in the limit of a weak feedback ($\alpha \ll 1$) we can determine the stability of the LO from looking at the phase error $\langle\delta\Phi_n^2\rangle^{1/2}$ for each Ramsay sequence independently. We shall therefore now explore the limitations arising from the noise term in this expression.

Equation (20) shows three kinds of noises that limit the stability. The first term is due to the uncertainty in $\hat{J}_z$ (ΔJ_z) and is proportional to $\delta\Phi_{n-1}$. This is the noise that limits the performance of the conventional protocol in Ref. [5] and the idea of the adaptive protocol is to eliminate this noise by minimizing $\delta\Phi_{n-1}$ through many measurements and feedback. The second term is the noise in $\hat{J}_y$, which is the term we want to decrease by squeezing the atomic spin. The final term is the accumulated back action of the measurements. The weaker the measurements are the smaller the contribution from this term will be.

We now analyze the various terms in $\langle\delta\Phi_n^2\rangle$ in more detail. The estimated phase in the j'th weak measurement is

$$\delta\phi_j^e = \left(\sin(\delta\Phi_{j-1})\hat{J}_z + \cos(\delta\Phi_{j-1})\hat{J}_y + \delta\hat{J}_{3,j} - \hat{P}_j/\Omega_j\right)/\langle\hat{J}_z\rangle. \tag{21}$$

Using equation (15)-(17) and equation (20)-(21) we can express the dominant contributions to the stability due to the noise in $\hat{J}_z$ as

$$\langle\delta\phi_0^2\rangle\langle(1 - \hat{J}_z/\langle\hat{J}_z\rangle)^{2n}\rangle \tag{22}$$

$$+2\langle\delta\phi_0(\sin(\delta\phi_0) - \delta\phi_0)\rangle\langle(1 - \hat{J}_z/\langle\hat{J}_z\rangle)^{2n-1}\hat{J}_z/\langle\hat{J}_z\rangle\rangle \tag{23}$$

$$+\langle(\sin(\delta\phi_0) - \delta\phi_0)^2\rangle\langle(1 - \hat{J}_z/\langle\hat{J}_z\rangle)^{2n-2}\hat{J}_z^2/\langle\hat{J}_z\rangle^2\rangle. \tag{24}$$

Here "dominant" refers to decreasing slowest with N. The above expressions are independent of the atomic state but we will now focus on a specific type of states in order to treat the system in more detail.

We consider spin squeezed states of the form $|\psi(\kappa)\rangle = \mathcal{N}(\kappa)\sum_m(-1)^m e^{-(m/\kappa)^2}|m\rangle$, where $|m\rangle$ are eigenstates of $\hat{J}_y$ with eigenvalue m, $\mathcal{N}(\kappa)$ is a normalization constant and the sum is from $-J$ to J where $J = N/2$ is the total angular momentum quantum number. κ gives the width of the state and thus determines the degree of squeezing. Uncorrelated atoms are well approximated by $\kappa = N^{1/2}$ while highly squeezed states are obtained for $\kappa \to 1$. We consider the limit where $N \gg 1$ such that we can replace sums with integrals when calculating the moments of the angular momentum operators. This allows us to get analytical expressions for the moments of J_z in equation (22)-(24). All the terms will decrease with growing κ since ΔJ_z decreases for growing κ, e.g., for a coherent spin state

$\kappa = \sqrt{N}$ we have $\Delta J_z \sim 0$. For a fixed κ all three terms will also decrease with growing n until a certain $n_{max}(\kappa)$ is reached. At this point the uncertainty in $\hat{J}_z$ prevents us from information gained in further measurements. We believe that this effect is due to the finite probability of measuring a J_z with opposite sign than $\langle\hat{J}_z\rangle$, which spoils the measurement strategy. Note that $n_{max}(\kappa)$ grows with κ since the width of $\hat{J}_z$ decreases with κ.

Of the three therm in Eqs. (22)-(24), the term in equation (24) is decreasing the slowest with n since this contains $(1-\hat{J}_z/\langle\hat{J}_z\rangle)$ to the lowest power. In the following we therefore focus on this term. In order for the performance to be nearly Heisenberg limited we need the contribution from this term to be close to or smaller than then Heisenberg limit, and this put restrictions on the possible values of κ. To determine the conditions for κ we have numerically solved the equation $\langle(1-\hat{J}_z/\langle\hat{J}_z\rangle)^{2n-2}\hat{J}_z^2/\langle\hat{J}_z\rangle^2\rangle = 1/N^2$ since $1/N^2$ is the Heisenberg limit of $\langle\delta\Phi_n^2\rangle$. This condition thus determines the parameters for which the noise from ΔJ_z is comparable to the Heisenberg limit. For N in the range $N = 10^3 - 10^9$ we have identified the n for which the equation is fulfilled for the smallest κ. The result is that in order for this term not to dominate the noise, κ needs to grow with increasing N to suppress the noise in J_z, but the growth can be slower than $\kappa \sim \log\sqrt{N}+2$ and requires $n \sim 3\log N$ measurements. Thus if we choose $\kappa \sim \log\sqrt{N}+2$ and $n \sim 3\log N$ the noise terms in equation (22)-(24) will decrease as $\lesssim 1/N^2$, which is the Heisenberg limit.

We now turn to the measurement noise, which consists of two parts. One is the accumulated back action contained in $\langle\delta\hat{J}_{3,n}^2\rangle$ (see equation (20)) while the other is due to the noise in the probe light (see the last term in equation (21)). For now we consider the accumulated back action. Using equation (15)-(17) we find that this is dominated by $\sum_{i=1}^{n-1}\Omega_i^2\langle\hat{X}_i^2\rangle\langle(\delta\phi_i^e\hat{J}_x/\langle\hat{J}_z\rangle)^2\rangle$ and that the dominating terms from each measurement are

$$i = 1: \quad \frac{1}{2}\langle\sin(\delta\phi_0)^2\rangle\langle(1-\hat{J}_z/\langle\hat{J}_z\rangle)^2\hat{J}_x^2/\langle\hat{J}_z\rangle^2\rangle\Omega_1^2 \tag{25}$$

$$i > 1: \quad \frac{1}{2}\langle(\sin(\delta\phi_0)-\delta\phi_0)^2\rangle\langle(1-\hat{J}_z/\langle\hat{J}_z\rangle)^{2i-2}\hat{J}_z^2\hat{J}_x^2/\langle\hat{J}_z\rangle^4\rangle\Omega_i^2 \tag{26}$$

where we have assumed that the probe light has vacuum statistics such that $\langle\hat{P}\rangle = \langle\hat{X}\rangle = \langle\hat{X}\hat{P}\rangle = 0$ and $\langle\hat{X}^2\rangle = \langle\hat{P}^2\rangle = 1/2$. Again we can get analytical expressions for $\langle(1-\hat{J}_z/\langle\hat{J}_z\rangle)^2\hat{J}_x^2/\langle\hat{J}_z\rangle^2\rangle$ and $\langle(1-\hat{J}_z/\langle\hat{J}_z\rangle)^{2i-2}\hat{J}_z^2\hat{J}_x^2/\langle\hat{J}_z\rangle^4\rangle$ by using the Gaussian approximation. Motivated by the previous numerical calculations we set $\kappa = \log\sqrt{N}+2$, $n \sim 3\log N$ and by numerically evaluating the terms for $N = 10^3 \rightarrow 10^9$ we find that for $\Omega_i = N^{-1+\frac{i}{n+1}}$ all terms will be $\lesssim 1/N^2$, i.e. at the Heisenberg limit.

We now consider the part of the measurement noise that comes from the noise in the probe light. From equation (15)-(17) and equation (21) we find that the dominating terms are $\sum_{i=1}^{n-1}\langle(1-\hat{J}_z/\langle\hat{J}_z\rangle)^{2n-2i}\rangle\langle\hat{P}_i^2\rangle/(\langle\hat{J}_z\rangle^2\Omega_i^2)$. Again by numerically evaluating the terms for $N = 10^3 \rightarrow 10^9$ we get the scaling of the terms and we find that for $\kappa = \log\sqrt{N}+2$, $n \sim 3\log N$ and $\Omega_i = N^{-1+\frac{i}{n+1}}$ all the terms will be $\lesssim 1/N^2$.

So far we have found that we can make the noise from the measurements and from ΔJ_z be $\lesssim 1/N^2$, which is the Heisenberg limit. The limiting noise in $\langle\Phi_n^2\rangle$ is then the noise from ΔJ_y. From (20) we find that this has a contribution of $\langle\hat{J}_y^2\rangle/\langle\hat{J}_z\rangle^2 = \Delta\hat{J}_y/\langle\hat{J}_z\rangle^2$. For the states $|\psi(\kappa)\rangle$ we find that $\Delta\hat{J}_y/\langle\hat{J}_z\rangle^2 \sim \kappa^2/N^2$. We thus get $\sigma_\gamma \sim (2/N+\log\sqrt{N}/N)/\sqrt{\gamma T}$ for $\kappa = \log\sqrt{N}+2$, $n \sim 3\log N$ and $\Omega_i = N^{-1+\frac{i}{n+1}}$. Note that we are in the limit of $N \gg 1$ and that σ_γ is in units of $(\gamma/(\omega^2\tau))^{1/2}$ (see article). For comparison the Heisenberg limit for the same Ramsey time is $\sigma_\gamma = (1/N)/\sqrt{\gamma T}$ in the same units. Hence our results shows that near Heisenberg limited stability can be obtained with the adaptive protocol.

We will also consider the performance of the adaptive protocol with uncorrelated atoms ($\kappa = \sqrt{N}$) . In this case as for the conventional Ramsey protocol the stability will be limited by the projection noise and will be $\sigma_\gamma \sim N^{-1/2}/\sqrt{\gamma T}$. Furthermore we will compare with the optimized protocol of Ref.[5] where the stability is $\sigma_\gamma \sim N^{-2/3}/\sqrt{\gamma T}$.

NUMERICAL SIMULATION

To verify the above findings we have simulated an atomic clock with a LO subject to both white and $1/f$ noise. For both types of noise we have simulated the clock for atom numbers ranging from 100 to 10^6. For $N \leq 1000$ we simulate the full quantum evolution during the measurements by bringing the input state $|\psi(\kappa)\rangle$ trough a Ramsey sequence and mixing it with a light state, which is assumed to be in vacuum. We pick the measurement outcome of $\hat{P}'$ according to the corresponding probability distribution and subsequently update the state of the atoms for the next measurement etc. We denote this as 'full quantum simulation'. For $N > 1000$ we approximate the probability distributions of $\hat{J}_{x,y,z}$ with Gaussian distributions with moments calculated from $|\psi(\kappa)\rangle$ in the limit of $N \gg 1$ such that we can replace the sum over m with an integral. We denote this as 'Gaussian simulation'.

The adaptive measurement protocol and the feedback on the LO are simulated as described in equation (1) and below in our article. The feedback effectively locks the LO to the atoms, and lowers the noise level of the LO as shown in Fig. 4 where the noise spectrum $S(f)$ of the LO is plotted against the frequency f. The noise spectrum is defined as $S(f)\delta(f+f') = \langle \delta\omega(f)\delta\omega(f')\rangle$ where $\delta\omega(f)$ is the Fourier transform of the frequency fluctuations $\delta\omega(t)$ of the LO. For a free running LO with white noise we use $S(f) = \gamma$ while for 1/f noise we use $S(f) = \gamma^2/f$ where γ is a parameter characterizing the fluctuations of the LO. The figure show that for high frequencies the locked LO has the noise of the free running oscillator but for low frequencies the LO is locked to the atoms and is limited by the atomic noise. Furthermore Fig. 4 shows how squeezing improves the stability of the clock by lowering the noise level of the locked LO more than for uncorrelated atoms. While the conventional Ramsey scheme works ideally for $\kappa \sim 14$, the adaptive protocol allows for $\kappa \sim 3$ at the atom number $N = 1000$ used in the figure, and thus leads to an improved stability.

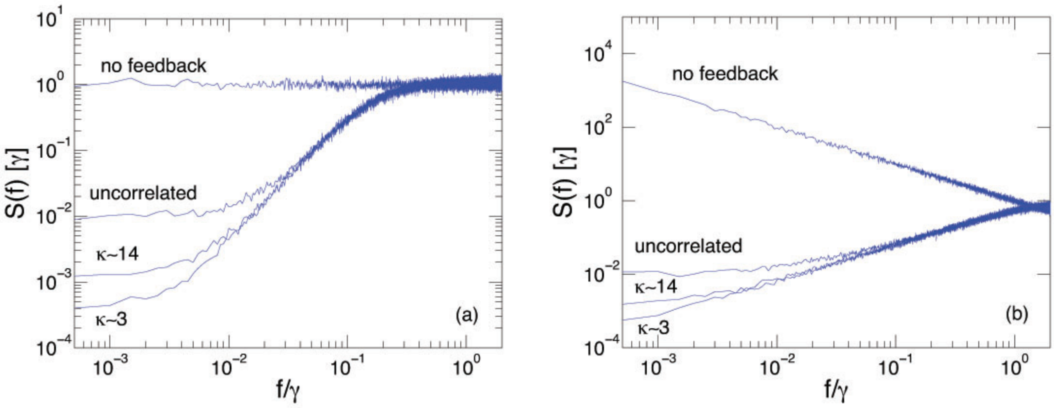

FIG. 4. (Colour online) Noise spectrum of the slaved LO for (a) white noise and (b) $1/f$-noise. The noise spectrum is defined as $S(f)\delta(f+f') = \langle \delta\omega(f)\delta\omega(f')\rangle$ where $\delta\omega(f)$ is the Fourier transform of the frequency fluctuations $\delta\omega(t)$ of the LO. The plots show how the feedback effectively locks the LO to the atoms and the noise in the LO becomes limited by the atomic noise for low frequencies. Furthermore it is seen how squeezing lowers the atomic noise and hence the noise of the locked LO. The plots where made for $N = 1000$ and $\gamma T = 0.1$ where the optimal performance of the conventional protocol is reached for a squeezing of $\kappa \sim 14$. The adaptive protocol allows for $\kappa \sim 3$ which further reduces the noise of the slaved LO. Note that the optimal performance of the adaptive protocol is reached at higher γT.

For a fixed Ramsey time the feedback strength (α) determines how long time the clock has to run before the LO is locked to the atoms. Since we simulate a clock running for a long but finite time there will be some remaining information from the last measurement results, which have not been fully exploited by the feedback loop. In our simulations we therefore do an additional phase correction to the LO after the final measurement. In principle the influence of the last few measurements could also have been reduced by running the simulation for a longer time, but by doing the correction we reduce the required simulation time. To find the required phase correction and obtain an expression for the stability of the clock we study the phase of the locked LO. At time $t_k = kT$ the phase of the LO is

$$\delta\phi(t_k) = \int_{t_{k-1}}^{t_k} \left(\delta\omega_0(t) + \sum_{i=1}^{k-1} \Delta\omega_i \right) \mathrm{d}t, \tag{27}$$

where $\delta\omega_0(t)$ is the frequency fluctuations of the unlocked LO and $\Delta\omega_i$ is the frequency corrections applied at time t_i. Using that $\Delta\omega_i = -\alpha\delta\phi_e(t_i)/T$ where $\delta\phi_e(t_i)$ is the estimated phase of the LO at time t_i we can write

$$\delta\phi(t_k) = \delta\phi_0(t_k) - \alpha \sum_{i=1}^{k+1} \delta\phi_e(t_i), \tag{28}$$

where $\delta\phi_0(t_k) = \int_{t_{k-1}}^{t_k} \delta\omega_0(t)\mathrm{d}t$. The mean frequency offset of the LO after running for a period $\tau = lT$ ($l \gg 1$) is

$$\delta\bar{\omega}(\tau) = \frac{1}{\tau} \left(\sum_{i=1}^{l} \delta\phi(t_i) - \phi_{\text{final correct}} \right), \tag{29}$$

where $\phi_{\text{final correct}}$ is the phase correction that we apply after the final measurement. Combining equation (28) and (29) we find that $\phi_{\text{final correct}} = \sum_{i=1}^{l} \left((1-\alpha)^{l-i} \delta\phi_e(t_i) + \sum_{j=1}^{i-1} \alpha(1-\alpha)^{l-i} \delta\phi_e(t_j) \right)$ will give the ideal performance. For this $\phi_{\text{final correct}}$ the mean frequency offset becomes

$$\delta\bar{\omega}(\tau) = \frac{1}{l} \sum_{i=1}^{l} \frac{\delta\phi(t_i) - \delta\phi_e(t_i)}{T}. \tag{30}$$

We use this expression to determine the stability of the clock, which is given by $\sigma_\gamma(\tau) = \langle (\delta\bar{\omega}(\tau)/\omega)^2 \rangle^{1/2}$. Note that while the sum is over different time intervals we cannot in general determine the stability by looking at an interval independently since the phases are correlated for finite α or for correlated noise in the free running LO e.g. $1/f$ noise. In our analytical calculations, however, we assume white noise and $\alpha \to 0$ so that we can ignore the correlations and consider each Ramsay sequence independently.

For the model investigated here, the stability increases with the Ramsay time but T is limited by two types of errors. For experiments or simulations running with a fixed Ramsay time there will always be a finite probability that the feedback loop jumps to a state with a phase difference of 2π (so called *fringe hops* [6]) or that a phase jump that is large enough to spoil the measurement strategy occurs. For the adaptive scheme this happens for phase jumps $\gtrsim \pi$ while it happens for phase jumps $\gtrsim \pi/2$ for the conventional protocol. The reason for this is that the adaptive protocol is able to distinguish whether the phase lies in the intervals $[0; \pi/2]$ or $[\pi/2; \pi]$ since they will lead to different response when we rotate the state during the feedback. On the contrary the conventional protocol only has a single projective measurement and cannot distinguish in which of the two intervals the phase lies. In our simulations we see the phase jumps as an abrupt break down as we increase the width of the distribution of the acquired phase, e.g. $\sigma^2 = \gamma T$ for white noise. This break down is clearly visible in Fig. 2 (reproduced as part (a) of Fig 5 where we also show the similar plot for $1/f$ noise (b)). Below we will investigate this breakdown in more detail.

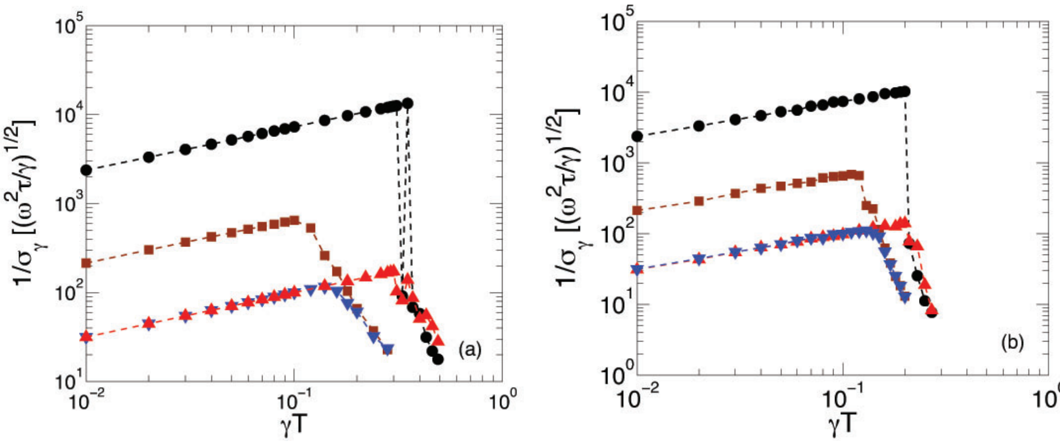

FIG. 5. Stability as a function of the Ramsey time (γT) for white noise (a) and $1/f$ (b) noise in the LO. The plots were made with $N = 10^5$. $\square, \triangledown$ are the best non-linear protocol of Ref. [5] while $\circ, \triangle$ are the adaptive protocol discussed in this article. The adaptive protocol allows for $\gamma T \sim 0.3$ and 0.2 for white and $1/f$ - noise while the conventional protocol of Ref. [5] only allows for $\gamma T \sim 0.1$ for both white and $1/f$ noise. $\triangledown, \triangle$ correspond to uncorrelated atoms while $\circ, \square$ are the ideal choices of squeezing in the adaptive and conventional protocols respectively. The probabilistic nature of the errors that limits γT is seen in (a) where the stability for the adaptive protocol jumps back and forth for $\gamma T > 0.3$ before it is definitely diminished.

For a simulation running for a time $\tau = lT$, we have l samples of the accumulated phase of the LO during the Ramsey time T. Assuming that these samples have an independent Gaussian probability distribution with zero mean (this will be the case for white noise and is approximately true for $1/f$ noise when the LO is locked to the atoms), the probability of all of these phases to be less than a critical value (a) is

$$P(\leq a) = \left(1 - \text{erfc}\left(\frac{a}{\sqrt{2}\sigma}\right)\right)^l \tag{31}$$

where the variance of the distribution σ depends on γT. For large l and as a function of σ this probability will drop abruptly from ~ 1 to ~ 0 around a certain $\sigma = \sigma_{max}$. Defining σ_{max} to be the position where $P(\leq a) = 1/2$ we find

that

$$1/2 = \left(1 - \mathrm{erfc}\left(\frac{a}{\sqrt{2}\sigma_{max}}\right)\right)^l. \tag{32}$$

Solving this equation gives

$$\sigma_{max} \approx \frac{a}{\sqrt{\ln(2/\pi) + 2\ln(l) - 2\ln(\ln 2) - \ln(\ln(2/\pi) + 2\ln(l) - 2\ln(\ln 2))}}, \tag{33}$$

where we have expanded to first order in $z = (1 - 2^{-1/l}) \sim \ln(2)/l$. It is seen that the breakdown (σ_{max}) has a weak (logarithmic) dependence on l. Solving equation (32) with the lhs. being equal to 0.95 and 0.05 we find that $P(\leq a)$ drops from 0.95 to 0.05 within a window of $\sim 2\sigma_{max}/\ln(l)$. Hence for large l the errors will appear very abruptly in the simulations.

Ideally we should include correction strategies for the errors due to large phase jumps in our simulations (e.g. running with different Ramsey times would correct for fringe hops), but for simplicity we ignore this. This means that our simulations have a weak dependence on the number of steps we simulate, but since it is only a logarithmic correction we do not expect this to change our results significantly. Instead we find the upper limit of γT from the simulations plotted in Fig. 2, where $l = 10^6$.

When minimizing the stability for the adaptive protocol we numerically minimize in both the degree of squeezing, the number of weak measurements and the strengths of the measurements. For white noise we can use the expression for σ_γ in the limit of $\alpha \ll 1$. For $1/f$ noise we include the feedback on the LO as described in equation (2) and below in the article. Note that $0 < \alpha < 1$ in order for the feedback to stabilize the LO [25]. We have used $\alpha = 0.1$ in our simulations and do not expect our results to change significantly for a different choice of α. We also minimize the stability in the degree of squeezing for the conventional protocol of Ref. [5] for comparison with the adaptive protocol.